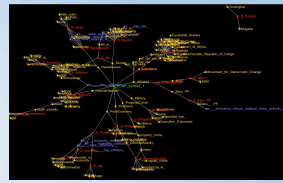


# Indago: Digital Knowledge Discovery

*The Power of Contextual Analysis*



***20:1 reduction of content allows analysis of millions of entries (e.g., cataloging of bibliographic records, email traffic analysis, etc. ) in real time***

## Co-Designed Solution

- Co-designed hardware solution achieves significantly accelerated content reduction, search, and contextual analysis
- NetLogic NLS220HAP cards, rated at 10-20Gbps, perform deep file inspection.
- Multi-engine parallel processing using lightweight threads.
- Highly scalable design that allows 5 orders of magnitude improvement over software-only pattern-matching equivalent processing.



- The most common and simplest form of analyzing a large repository uses simple word-occurrence searching techniques in which a set of terms are used to search for relevant content
- A significant drawback of this approach is that the results contain terms in the wrong context—for example, the term "Apple" gets results related to fruits and computers
- LANL's Digital Knowledge Discovery (DKD) team is creating complex models of language contained in unstructured text to build increasingly sophisticated tools that move beyond word occurrence to contextual analysis, giving us the ability to discover patterns found in large collections

Jorge Roman, [jhr@lanl.gov](mailto:jhr@lanl.gov) and phone  
Dave DuBois, [dhd@lanl.gov](mailto:dhd@lanl.gov) and phone